



16001
Raw Seq 16001
7/31/03

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/647,522

DATE: 03/18/2002
TIME: 15:20:39

Input Set : A:\Nagai.h.app
Output Set: N:\CRF3\03182002\I647522.raw

3 <110> APPLICANT: NAGAI, Hiroshi
4 NAKAJIMA, Terumi
6 <120> TITLE OF INVENTION: NOVEL HEMOLYTIC ACTIVE PROTEINS AND GENES ENCODING THE
7 SAME
9 <130> FILE REFERENCE: PCTJP9901607NAGAIHiroshietal 183049264
11 <140> CURRENT APPLICATION NUMBER: US/09/647,522
12 <141> CURRENT FILING DATE: 2000-10-02
14 <150> PRIOR APPLICATION NUMBER: PCT/JP99/01607
15 <151> PRIOR FILING DATE: 1999-03-30
17 <150> PRIOR APPLICATION NUMBER: JP 10/88569
18 <151> PRIOR FILING DATE: 1998-04-01
20 <160> NUMBER OF SEQ ID NOS: 18
22 <170> SOFTWARE: PatentIn Ver. 2.0
24 <210> SEQ ID NO: 1
25 <211> LENGTH: 14
26 <212> TYPE: PRT
27 <213> ORGANISM: Carybdea rastonii
29 <220> FEATURE:
30 <223> OTHER INFORMATION: This amino acid residue sequence corresponds to
31 amino acid residue positions 56-69 of SEQ ID NO:5.
33 <400> SEQUENCE: 1
34 Gly Glu Ile Gln Thr Lys Pro Asp Arg Val Gly Gln Ala Thr
35 1 5 10
38 <210> SEQ ID NO: 2
39 <211> LENGTH: 18
40 <212> TYPE: PRT
41 <213> ORGANISM: Carybdea rastonii
43 <220> FEATURE:
44 <223> OTHER INFORMATION: This amino acid residue sequence corresponds to
45 amino acid residue positions 250-267 of SEQ ID
46 NO:5.
48 <400> SEQUENCE: 2
49 Gly Asn Ala Glu His Val Ala Ser Ala Val Glu Asn Ala Asn Arg Val
50 1 5 10 15
52 Asn Lys
56 <210> SEQ ID NO: 3
57 <211> LENGTH: 15
58 <212> TYPE: PRT
59 <213> ORGANISM: Carybdea rastonii
61 <220> FEATURE:
62 <223> OTHER INFORMATION: This amino acid residue sequence corresponds to
63 amino acid residue positions 363-377 of SEQ ID
64 NO:5.

ENTERED

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/647,522

DATE: 03/18/2002
TIME: 15:20:39

Input Set : A:\Nagai.h.app
Output Set: N:\CRF3\03182002\I647522.raw

66 <400> SEQUENCE: 3
67 Met Ser Asp Gly Phe Tyr Thr Met Glu Asn Ser Asp Arg Arg Lys
68 1 5 10 15
71 <210> SEQ ID NO: 4
72 <211> LENGTH: 1610
73 <212> TYPE: DNA
74 <213> ORGANISM: Carybdea rastonii
76 <220> FEATURE:
77 <221> NAME/KEY: CDS
78 <222> LOCATION: (28)..(1380)
80 <220> FEATURE:
81 <221> NAME/KEY: protein_bind
82 <222> LOCATION: (1381)..(1610)
84 <400> SEQUENCE: 4
85 gcacaaggcga cttggtaag gagcacc atg att ctg aaa cat ctt cct tgg ctc 54
86 Met Ile Leu Lys His Leu Pro Trp Leu
87 1 5
89 ttt att gtc ctt gca att act tct gca aaa cat ggc aaa cgc tct gat 102
90 Phe Ile Val Leu Ala Ile Thr Ser Ala Lys His Gly Lys Arg Ser Asp
91 10 15 20 25
93 gtc aat tct tta ctt act aag gta gaa act gcc tta aaa gaa gct tct 150
94 Val Asn Ser Leu Leu Thr Lys Val Glu Thr Ala Leu Lys Glu Ala Ser
95 30 35 40
97 ggt agc aac gag gct gct ctt gag gct tta gag ggc tta aaa gga gag 198
98 Gly Ser Asn Glu Ala Ala Leu Glu Ala Leu Glu Gly Leu Lys Gly Glu
99 45 50 55
101 atc cag aca aaa cca gac cga gtt gga caa gcc aca aaa atc ctt gga 246
102 Ile Gln Thr Lys Pro Asp Arg Val Gly Gln Ala Thr Lys Ile Leu Gly
103 60 65 70
105 tct gtc gga tca gct cta gga aaa tta aat tct gga gat gca acc aaa 294
106 Ser Val Gly Ser Ala Leu Gly Lys Leu Asn Ser Gly Asp Ala Thr Lys
107 75 80 85
109 atc att tct ggt tgc ctc gac att gtt gca gga att gca aca act ttt 342
110 Ile Ile Ser Gly Cys Leu Asp Ile Val Ala Gly Ile Ala Thr Thr Phe
111 90 95 100 105
113 gga ggc cct gtc ggg atg gga atc gga gcc gta gct tct ttt gtt tct 390
114 Gly Gly Pro Val Gly Met Gly Ile Gly Ala Val Ala Ser Phe Val Ser
115 110 115 120
117 tca att cta tca ttg ttt act gga agc tca gca aag aac tca gtt gct 438
118 Ser Ile Leu Ser Leu Phe Thr Gly Ser Ser Ala Lys Asn Ser Val Ala
119 125 130 135
121 gcc gtt att gat aga gct tta agc aag cat cgc gat gag gcc atc caa 486
122 Ala Val Ile Asp Arg Ala Leu Ser Lys His Arg Asp Glu Ala Ile Gln
123 140 145 150
125 aga cat gca gca ggt gcc aag aga gat ttt gct gaa tca tct gca ttc 534
126 Arg His Ala Ala Gly Ala Lys Arg Asp Phe Ala Glu Ser Ser Ala Phe
127 155 160 165
129 att cag gtc atg aaa cag cag tcc aat ctt aca gat agc gac cta agt 582
130 Ile Gln Val Met Lys Gln Gln Ser Asn Leu Thr Asp Ser Asp Leu Ser

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/647,522

DATE: 03/18/2002
TIME: 15:20:39

Input Set : A:\Nagai h.app
Output Set: N:\CRF3\03182002\I647522.raw

131	170	175	180	185													
133	atc	att	gca	gct	aat	gtt	cct	gtt	tat	aaa	ttt	agt	aat	ttt	atc	gga	630
134	Ile	Ile	Ala	Ala	Asn	Val	Pro	Val	Tyr	Lys	Phe	Ser	Asn	Phe	Ile	Gly	
135																	200
137	cag	ttg	gag	agc	aga	att	tcc	caa	ggc	gca	gca	act	acc	agt	ctt	agc	678
138	Gln	Leu	Glu	Ser	Arg	Ile	Ser	Gln	Gly	Ala	Ala	Thr	Thr	Ser	Leu	Ser	
139																	
141	gat	gca	aag	aga	gcc	gtt	gac	ttc	att	ctg	ctc	tat	tgt	caa	ctt	gta	726
142	Asp	Ala	Lys	Arg	Ala	Val	Asp	Phe	Ile	Leu	Leu	Tyr	Cys	Gln	Leu	Val	
143																	
145	gtc	atg	aga	gaa	acc	ttg	ctg	gtc	gac	ttg	gct	att	ctc	tac	agg	aaa	774
146	Val	Met	Arg	Glu	Thr	Leu	Leu	Val	Asp	Leu	Ala	Ile	Leu	Tyr	Arg	Lys	
147																	
149	gga	aat	gca	gaa	cac	gtg	gca	agt	gct	gtg	gaa	aac	gct	aat	agg	gta	822
150	Gly	Asn	Ala	Glu	His	Val	Ala	Ser	Ala	Val	Glu	Asn	Ala	Asn	Arg	Val	
151																	
153	aac	aaa	gag	cta	gct	gat	acc	cta	gat	ttt	ctt	cat	aaa	ttg	att	870	
154	Asn	Lys	Glu	Leu	Ala	Ala	Asp	Thr	Leu	Asp	Phe	Leu	His	Lys	Leu	Ile	
155																	
157	cct	gaa	caa	gca	ttg	ata	ggt	gca	gtt	tat	cat	cca	att	tct	gcc	tct	918
158	Pro	Glu	Gln	Ala	Leu	Ile	Gly	Ala	Val	Tyr	His	Pro	Ile	Ser	Ala	Ser	
159																	
161	gaa	act	agc	aaa	gca	ata	tta	aat	tac	acg	aaa	tac	ttt	gga	gtt	cca	966
162	Glu	Thr	Ser	Lys	Ala	Ile	Leu	Asn	Tyr	Thr	Lys	Tyr	Phe	Gly	Val	Pro	
163																	
165	gat	gtt	ccc	cgt	cct	att	gga	aac	cgc	aga	tac	aaa	ttt	aca	aat	agt	1014
166	Asp	Val	Pro	Arg	Pro	Ile	Gly	Asn	Arg	Arg	Tyr	Lys	Phe	Thr	Asn	Ser	
167																	
169	tac	tgg	aat	acc	tac	agt	ata	tgc	agt	gag	gct	tac	atg	gga	aat	tac	1062
170	Tyr	Trp	Asn	Thr	Tyr	Ser	Ile	Cys	Ser	Glu	Ala	Tyr	Met	Gly	Asn	Tyr	
171																	
173	atg	ttc	aga	ggc	tgt	tct	aac	gtt	cg	aat	cca	aat	atc	agg	gta	tcc	1110
174	Met	Phe	Arg	Gly	Cys	Ser	Asn	Val	Arg	Asn	Pro	Asn	Ile	Arg	Val	Ser	
175																	
177	aaa	atg	tct	gat	ggg	ttt	tac	acc	atg	gag	aat	agc	gat	cg	agg	aag	1158
178	Lys	Met	Ser	Asp	Gly	Phe	Tyr	Thr	Met	Glu	Asn	Ser	Asp	Arg	Arg	Lys	
179																	
181	ttg	tat	atc	acc	aag	cat	gac	caa	gga	tgg	gga	tgg	ggt	act	ttg	gat	1206
182	Leu	Tyr	Ile	Thr	Lys	His	Asp	Gln	Gly	Trp	Gly	Trp	Gly	Thr	Leu	Asp	
183																	
185	gag	gat	cca	gg	tgt	gac	caa	ggc	cat	atg	agg	ttc	att	cct	ttg	aga	1254
186	Glu	Asp	Pro	Gly	Asp	Gln	Gly	His	Met	Arg	Phe	Ile	Pro	Leu	Arg	His	
187																	
189	ggg	aag	tat	atg	gta	agc	tct	aag	agg	tgg	ccc	aac	tgg	ttc	atg	tat	1302
190	Gly	Lys	Tyr	Met	Val	Ser	Ser	Lys	Arg	Trp	Pro	Asn	Trp	Phe	Met	Tyr	
191																	
193	atg	gaa	tca	agt	gcc	agt	ggc	tac	att	cgc	agc	tgg	gaa	aat	aat	cca	1350
194	Met	Glu	Ser	Ser	Ala	Ser	Gly	Tyr	Ile	Arg	Ser	Trp	Glu	Asn	Asn	Pro	
195																	

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/647,522

DATE: 03/18/2002
TIME: 15:20:39

Input Set : A:\Nagai.h.app
Output Set: N:\CRF3\03182002\I647522.raw

197 gga cct caa gga cat tgg agt ata aca taa ttaaaagagga atcaacaatg 1400
 198 Gly Pro Gln Gly His Trp Ser Ile Thr
 199 445 450
 201 tcccaaaggc atacgaatat aagacatcaa acgaatgcag tacttaaagt gcacacttgt 1460
 203 atttctacat aggatgtcgt catgaaagt cataaaccat ccagcgact aatttcatat 1520
 205 taaacattaa tgttcctta taatgcattt tcattgaaatc tctattgtga catttcaaga 1580
 207 ggatatgttt gaaagaaaaca aaaaaaaaaa 1610
 210 <210> SEQ ID NO: 5
 211 <211> LENGTH: 450
 212 <212> TYPE: PRT
 213 <213> ORGANISM: Carybdea rastonii
 215 <400> SEQUENCE: 5
 216 Met Ile Leu Lys His Leu Pro Trp Leu Phe Ile Val Leu Ala Ile Thr
 217 1 5 10 15
 219 Ser Ala Lys His Gly Lys Arg Ser Asp Val Asn Ser Leu Leu Thr Lys
 220 20 25 30
 222 Val Glu Thr Ala Leu Lys Glu Ala Ser Gly Ser Asn Glu Ala Ala Leu
 223 35 40 45
 225 Glu Ala Leu Glu Gly Leu Lys Gly Glu Ile Gln Thr Lys Pro Asp Arg
 226 50 55 60
 228 Val Gly Gln Ala Thr Lys Ile Leu Gly Ser Val Gly Ser Ala Leu Gly
 229 65 70 75 80
 231 Lys Leu Asn Ser Gly Asp Ala Thr Lys Ile Ile Ser Gly Cys Leu Asp
 232 85 90 95
 234 Ile Val Ala Gly Ile Ala Thr Thr Phe Gly Gly Pro Val Gly Met Gly
 235 100 105 110
 237 Ile Gly Ala Val Ala Ser Phe Val Ser Ser Ile Leu Ser Leu Phe Thr
 238 115 120 125
 240 Gly Ser Ser Ala Lys Asn Ser Val Ala Ala Val Ile Asp Arg Ala Leu
 241 130 135 140
 243 Ser Lys His Arg Asp Glu Ala Ile Gln Arg His Ala Ala Gly Ala Lys
 244 145 150 155 160
 246 Arg Asp Phe Ala Glu Ser Ser Ala Phe Ile Gln Val Met Lys Gln Gln
 247 165 170 175
 249 Ser Asn Leu Thr Asp Ser Asp Leu Ser Ile Ile Ala Ala Asn Val Pro
 250 180 185 190
 252 Val Tyr Lys Phe Ser Asn Phe Ile Gly Gln Leu Glu Ser Arg Ile Ser
 253 195 200 205
 255 Gln Gly Ala Ala Thr Thr Ser Leu Ser Asp Ala Lys Arg Ala Val Asp
 256 210 215 220
 258 Phe Ile Leu Leu Tyr Cys Gln Leu Val Val Met Arg Glu Thr Leu Leu
 259 225 230 235 240
 261 Val Asp Leu Ala Ile Leu Tyr Arg Lys Gly Asn Ala Glu His Val Ala
 262 245 250 255
 264 Ser Ala Val Glu Asn Ala Asn Arg Val Asn Lys Glu Leu Ala Ala Asp
 265 260 265 270
 267 Thr Leu Asp Phe Leu His Lys Leu Ile Pro Glu Gln Ala Leu Ile Gly
 268 275 280 285
 270 Ala Val Tyr His Pro Ile Ser Ala Ser Glu Thr Ser Lys Ala Ile Leu

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/647,522

DATE: 03/18/2002
TIME: 15:20:39

Input Set : A:\Nagai h.app
Output Set: N:\CRF3\03182002\I647522.raw

271 290 295 300
273 Asn Tyr Thr Lys Tyr Phe Gly Val Pro Asp Val Pro Arg Pro Ile Gly
274 305 310 315 320
276 Asn Arg Arg Tyr Lys Phe Thr Asn Ser Tyr Trp Asn Thr Tyr Ser Ile
277 325 330 335
279 Cys Ser Glu Ala Tyr Met Gly Asn Tyr Met Phe Arg Gly Cys Ser Asn
280 340 345 350
282 Val Arg Asn Pro Asn Ile Arg Val Ser Lys Met Ser Asp Gly Phe Tyr
283 355 360 365
285 Thr Met Glu Asn Ser Asp Arg Arg Lys Leu Tyr Ile Thr Lys His Asp
286 370 375 380
288 Gln Gly Trp Gly Trp Gly Thr Leu Asp Glu Asp Pro Gly Asp Gln Gly
289 385 390 395 400
291 His Met Arg Phe Ile Pro Leu Arg His Gly Lys Tyr Met Val Ser Ser
292 405 410 415
294 Lys Arg Trp Pro Asn Trp Phe Met Tyr Met Glu Ser Ser Ala Ser Gly
295 420 425 430
297 Tyr Ile Arg Ser Trp Glu Asn Asn Pro Gly Pro Gln Gly His Trp Ser
298 435 440 445
300 Ile Thr
301 450
304 <210> SEQ ID NO: 6
305 <211> LENGTH: 19
306 <212> TYPE: DNA
307 <213> ORGANISM: Artificial Sequence
309 <220> FEATURE:
310 <223> OTHER INFORMATION: Description of Artificial Sequence: Degenerate PCR
311 primer, 7-F, used in the cloning of the partial
312 cDNA of the hemolytic active protein of Carybdea
313 rastonii
315 <220> FEATURE:
316 <221> NAME/KEY: unsure
317 <222> LOCATION: (12)
318 <223> OTHER INFORMATION: n = inosine
320 <220> FEATURE:
321 <221> NAME/KEY: unsure
322 <222> LOCATION: (18)
323 <223> OTHER INFORMATION: n = inosine
325 <400> SEQUENCE: 6
326 garathcara cnaarcng 19
329 <210> SEQ ID NO: 7
330 <211> LENGTH: 19
331 <212> TYPE: DNA
332 <213> ORGANISM: Artificial Sequence
334 <220> FEATURE:
335 <221> NAME/KEY: unsure
336 <222> LOCATION: (2)
337 <223> OTHER INFORMATION: n = inosine
339 <220> FEATURE:

Use of n and/or Xaa has been detected in the Sequence Listing.
Review the Sequence Listing to insure a corresponding
explanation is presented in the <220> to <223> fields of
each sequence using n or Xaa.

VERIFICATION SUMMARY
PATENT APPLICATION: US/09/647,522

DATE: 03/18/2002
TIME: 15:20:40

Input Set : A:\Nagai h.app
Output Set: N:\CRF3\03182002\I647522.raw

L:11 M:270 C: Current Application Number differs, Replaced Application Number
L:12 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:326 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:6
L:351 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7
L:381 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8
L:411 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:9
L:436 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:10
L:461 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:11
L:486 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:12
L:511 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:13